LEVEL II

AD A096093

PURDUE UNIVERSITY

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

DEPARTMENT OF STATISTICS

DIVISION OF MATHEMATICAL SCIENCES

81 3 09 105

DDC FILE COPY

NONPARAMETRIC PROCEDURES IN MULTIPLE DECISIONS*
(RANKING AND SELECTION PROCEDURES)

by

Shanti S. Gupta
Purdue University

and

Gary C. McDonald
General Motors Research Laboratories

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #80-13

May 1980

*The writing of this paper was supported in part by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University.

NONPARAMETRIC PROCEDURES IN MULTIPLE DECISIONS*
(RANKING AND SELECTION PROCEDURES)

Shanti S. Gupta
Purdue University

and

Gary C. McDonald
General Motors Research Laboratories

ABSTRACT

This article surveys statistical techniques which are *nonparametric* in nature and used in formal ranking and selection of populations. Such methods have been developed only within the last fifteen years and are usually based on rank scores and/or robust estimators (such as the Hodges-Lehmann estimator). The procedures surveyed are applicable to one-way classifications, two-way classifications, and *paired-comparison models*. Computational methods, useful inequalities, and appropriate numerical tables required to implement these techniques are identified and discussed. Asymptotic relative efficiencies of the nonparametric methods, compared to their parametric counterparts, are presented. Specific applications of these methods (such as traffic fatality rates) are mentioned and areas for further theoretical and computational research are identified.

1. Introduction to Selection and Ranking Procedures

A common problem faced by an experimenter is one of comparing several categories or populations. These may be, for example, different varieties of

*The writing of this paper was supported in part by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University.

For presentation at: János Bolyai Mathematical Society: Colloquium on Nonparametric Statistical Inference, Budapest, Hungary, June 23-28, 1980.

a grain, different competing manufacturing processes for an industrial product, or different drugs (treatments) for a specific disease. In other words, we have $k(\geq 2)$ populations and each population is characterized by the value of a parameter of interest θ, which may be, in the example of drugs, an appropriate measure of the effectiveness of a drug. The classical approach to this problem is to test the homogeneity (null) hypothesis H_0: $\theta_1 = \ldots = \theta_k$, where $\theta_1, \ldots, \theta_k$ are the values of the parameter for these populations. In the case of normal populations with means $\theta_1, \ldots, \theta_k$ and a common variance σ^2, the test can be carried out using the F-ratio of the analysis of variance.

The above classical approach is inadequate and does not answer a frequently encountered experimenter's question, namely, how to identify the best category? In fact, the method of least significant differences based on t-tests has been used in the past to detect differences between the average yields of different varieties and thereby choose the 'best' variety. But this method (and others related to it) is indirect and does not easily provide an overall probability of a correct selection. Also the multiple comparison techniques developed largely by Tukey and Scheffé arose from the desire to draw inference about the populations when the homogeneity hypothesis is rejected.

Selection and Ranking Procedures

The formulation of a k-sample problem as a multiple decision problem enables the experimenter to answer questions regarding the best category. The formulation of multiple decision procedures in the framework of selection and ranking procedures has been accomplished generally using either the indifference zone approach or the (random sized) subset selection approach. The former approach was introduced by Bechhofer (1954). Substantial

contribution to the early and subsequent developments in the subset selection theory has been made by Gupta starting from his work in 1956. For more details about the numerous contributions and the related bibliography, reference should be made to a recently published book by Gupta and Panchapakesan (1979). This monograph discusses both approaches.

A Brief Description of the Two Approaches

Bechhofer (1954) considered the problem of ranking k normal means. In order to explain the basic formulation, consider the problem of selecting the population with the largest mean from k normal populations with unknown means μ_i, $i = 1,\ldots,k$, and a common known variance σ^2. Let $\bar{x}_i$, $i = 1,\ldots,k$, denote the means of independent samples of size n from these populations. The 'natural' procedure (which can be shown to have optimum properties) will be to select the population that yields the largest $\bar{x}_i$. The experimenter would, of course, need a guarantee that this procedure will pick the population with the largest μ_i with a probability not less than a specified level P*. For the problem to be meaningful P* lies between 1/k and 1. Since we do not know the true configuration of the μ_i, we look for the least favorable configuration (LFC) for which the probability of a correct selection, P(CS), will be at least P*. Since the LFC is given by $\mu_1 = \ldots = \mu_k$, the probability guarantee cannot be met whatever be the sample size n.

A natural modification is to insist on the minimum probability guarantee whenever the best population is sufficiently superior to the next best. In other words, the experimenter specifies a positive constant Δ^* and requires that the P(CS) is at least P* whenever $\mu_{[k]}-\mu_{[k-1]} \geq \Delta^*$, where $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ denote the ordered means. Now the minimization of P(CS) is over the part Ω_{Δ^*} of the parameter space in which $\mu_{[k]} - \mu_{[k-1]} \geq \Delta^*$. The complement of Ω_{Δ^*} is

called the indifference zone for the obvious reason. The LFC in Ω_{Δ^*} is given by $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \Delta^*$. The problem then reduces to determing the minimum sample size required in order to have $P(CS) \geq P^*$ for the LFC.

Bechhofer's formulation can be generalized from that described above. His general ranking problem includes, for example, selection of the t best populations.

In the subset selection approach, the goal is to select a non-empty subset of the populations so as to include the best population. Here the size of the selected subset is random and is determined by the observations themselves. In the case of normal populations with unknown means $\mu_1, \ldots, \mu_k$, and a common variance σ^2, the rule proposed by Gupta (1956) selects the population that yields $\bar{x}_i$ if and only if $\bar{x}_i \geq \max_{1 \leq j \leq k} \bar{x}_j - \frac{d\sigma}{\sqrt{n}}$, where $d = d(k, P^*) > 0$ is determined so that the P(CS) is at least P*. Here a correct selection is selection of any subset that includes the population with the largest μ_i. Thus, the LFC is with regard to the whole parameter space Ω. Under this formulation, for given k and P* we determine d. The rule explicitly involves n. In general, the rule will involve a constant which depends on k, P*, and n. The performance of a subset selection procedure is studied by evaluating the expected subset size and its supremum over the parameter space Ω.

Nonparametric Techniques in Multiple Decision Theory

In the present paper, we describe selection and ranking (ordering) procedures which are nonparametric or distribution-free. Such procedures have the desirable property that their applicability is valid under relatively mild assumptions regarding the underlying population(s) from which the data are obtained. Although the importance of nonparametric methods as a

significant branch of modern statistics is recognized by statisticians, modern nonparametric techniques are usually restricted to hypothesis testing, point estimators, confidence intervals, and multiple comparison procedures. Other recent advances in nonparametric tests can be found in Hollander and Wolfe (1973) and Lehmann (1975). The development of nonparametric methods for multiple decision procedures is important in statistical research. The present paper deals with selection procedures with special emphasis on the subset selection approach related to the largest unknown parameter. Analogous procedures (with proper modifications) are available for the selection in terms of the smallest parameter.

In Section 2, we discuss procedures based on the ranks in the combined sample. Section 3 deals with bounds on the probability of a correct selection associated with the first two procedures $R_1(G)$ and $R_2(G)$ of Section 2. In Section 3, the exact and asymptotic distribution of the (appropriate) statistic based on rank sums is discussed. In Section 5, we provide comparisons between R_1 and R_3 and certain parametric procedures in terms of their asymptotic relative efficiencies. Selection procedures based on pairwise ranks are discussed briefly in Section 6. Section 7 deals with selection procedures based on vector ranks. In Section 8, procedures based on Hodges-Lehmann estimators are discussed.

2. Procedures Based on Combined Ranks.

Let $\pi_1,\ldots,\pi_k$ be $k(\geq 2)$ independent populations. The associated random variables X_{ij}, $j = 1,\ldots,n_i$, $i = 1,\ldots,k$, are assumed independent and to have a continuous distribution $F_{\theta_i}(x)$, where θ_i belong to some interval Θ on the real line. Suppose $F_\theta(x)$ is a stochastically increasing (SI) family of distributions, i.e. if θ_1 is less than θ_2, then $F_{\theta_1}(x)$ and $F_{\theta_2}(x)$ are distinct and $F_{\theta_2}(x) \leq F_{\theta_1}(x)$ for all x. Examples of such families of distributions are: (1) any location parameter family, i.e. $F_\theta(x) = F(x-\theta)$; (2) any scale parameter family, i.e. $F_\theta(x) = F(x/\theta)$, $x > 0$, $\theta > 0$; (3) any family of distribution functions whose densities possess the monotone likelihood ratio (or TP_2) property. Let R_{ij} denote the rank of the observation x_{ij} in the combined sample; i.e. if there are exactly r observations less than x_{ij} then $R_{ij} = r+1$. These ranks are well-defined with probability one, since the random variables are assumed to have a continous distribution. Let $Z(1) \leq Z(2) \leq \ldots \leq Z(N)$ denote an ordered sample of size $N = \sum_{i=1}^{k} n_i$ from any continuous distribution G, such that

$$-\infty < a(r) \equiv E[Z(r)|G] < \infty \quad (r = 1,\ldots,N).$$

With each of the random variables X_{ij} associate the number $a(R_{ij})$ and define

$$H_i = n_i^{-1} \sum_{j=1}^{n_i} a(R_{ij}) \quad (i = 1,\ldots,k). \tag{2.1}$$

Using the quantities H_i, Gupta and McDonald (1970) have defined procedures for selecting a subset of the k populations. Letting $\theta_{[i]}$ denote the ith smallest unknown parameter, we have

$$F_{\theta_{[1]}}(x) \geq F_{\theta_{[2]}}(x) \geq \ldots \geq F_{\theta_{[k]}}(x), \quad \forall x. \tag{2.2}$$

The population whose associated random variables have the distribution $F_{\theta_{[k]}}(x)$ will be called the best population. In case several populations possess the largest parameter value $\theta_{[k]}$, one of them is tagged at random and called the best. A 'Correct Selection' (CS) is said to occur if and only if the best population is included in the selected subset. In the usual subset selection problem one wishes to select a subset such that the probability is at least equal to a preassigned constant $P^*(1/k < P^* < 1)$ that the selected subset includes the best population. Mathematically, for a given selection rule R,

$$\inf_{\Omega} P(CS|R) \geq P^*, \tag{2.3}$$

where $$\Omega = \{\underline{\theta} = (\theta_1,\ldots\theta_k): \theta_i \in \Theta, \quad i = 1,2,\ldots,k\}. \tag{2.4}$$

The following three classes of selection procedures, which choose a subset of the k given populations, and which depend on the given distribution G, have been considered:

$$R_1(G): \text{ Select } \pi_i \text{ iff } H_i \geq \max_{1\leq j\leq k} H_j - d \quad (i = 1,\ldots,k,\ d > 0), \tag{2.5}$$

$$R_2(G): \text{ Select } \pi_i \text{ iff } H_i \geq c^{-1} \max_{1\leq j\leq k} H_j \quad (i = 1,\ldots,k,\ c \geq 1), \tag{2.6}$$

$$R_3(G): \text{ Select } \pi_i \text{ iff } H_i \geq D \quad (i = 1,\ldots,k,\ -\infty < D < \infty). \tag{2.7}$$

It should be noted that rules $R_1(G)$, $R_2(G)$, and $R_3(G)$ are equivalent if $k = 2$. The procedures $R_1(G)$ (and their randomized analogs) have been suggested by Bartlett and Govindarajulu (1968) for continuous distributions differing by a location parameter. The procedure $R_2(G)$ will be studied in this paper only for the case where $H_i > 0$ for all i. The constants d and c are usually chosen to be as small as possible, D as large as possible, while satisfying the

probability requirement (2.3). The number of populations included in the selected subset is a random variable which takes values 1 to k inclusive for rules $R_1(G)$ and $R_2(G)$. The subset chosen by rule $R_3(G)$, however, could possibly be empty. This aspect will be addressed further at the end of Section 3.

It has been shown by Gupta and McDonald that the infimum of $P(CS|R_i(G))$, $i = 1,2,3$, over Ω is attained for $\underline{\theta} \in \Omega_k = \{\underline{\theta}: \theta_{[k-1]} = \theta_{[k]}\}$. This shows that for $k = 2$ the infimum occurs at an equi-parameter configuration.

For $k \geq 3$ the least favorable configuration (LFC) is not given by the equi-parameter configuration for $R_1(G)$ and $R_2(G)$ as can be seen from the counterexamples of Rizvi and Woodworth (1970). The counterexample distribution is a mixture of two distinct uniform random variables and is established for P* near 1.

For the procedure $R_3(G)$ we can say more about the infimum of the probability of a correct selection. The LFC is given by the equi-parameter configuration and so

$$\inf_{\Omega} P(CS|R_3(G)) = \inf_{\Omega_0} P(CS|R_3(G)),$$

where $\Omega_0 = \{\underline{\theta} \in \Omega: \theta_{[1]} = \ldots = \theta_{[k]}\}$.

These selection rules are called distribution-free (or nonparametric) if the constants required for implementation are computed from $P(CS|R_i(G)) = P^*$ for $\underline{\theta} \in \Omega_0$. In this case the probability does not depend on the common parameter value and on the underlying distribution functions. The probability computation is based on a random assignment of rank scores.

3. Bounds on $P(CS|R_i(G))$, $i = 1,2$.

Since the exact LFC for the selection rules $R_1(G)$ and $R_2(G)$ is unknown for $k > 2$, it is useful to have bounds for the probabilities of correct selection. We will assume $n_i = n$, $i = 1,\ldots,k$. First consider rule $R_1(G)$. Since

$$(k-1)^{-1} \sum_{j=1}^{k-1} H_{(j)} \leq \max_{1 \leq j \leq k-1} H_{(j)} \leq n^{-1} \sum_{r=N-n+1}^{N} a(r), \tag{3.1}$$

and

$$\sum_{j=1}^{k} H_{(j)} = A/n,$$

where $\sum_{r=1}^{N} a(r) = A$, it follows that

$$\inf_{\Omega} P(H_{(k)} \geq v) \leq \inf_{\Omega} P(CS|R_1(G)) \leq \inf_{\Omega} P(H_{(k)} \geq u). \tag{3.2}$$

The quantities u and v are defined by

$$u \equiv u(d,k,n) = [A-nd(k-1)]/nk, \tag{3.3}$$

and

$$v = v(d,k,n) = n^{-1} \sum_{r=N-n+1}^{N} a(r)-d. \tag{3.4}$$

For the rule $R_2(G)$, we get a similar expression:

$$\inf_{\Omega} P(H_{(k)} \geq v') \leq \inf_{\Omega} P(CS|R_2(G)) \leq \inf_{\Omega} P(H_{(k)} \geq u'), \tag{3.5}$$

where

$$u' \equiv u'(d,k,n) = n^{-1} A[1+c(k-1)]^{-1} \tag{3.6}$$

and

$$v' \equiv v'(d,k,n) = (nc)^{-1} \sum_{r=N-n+1}^{N} a(r). \tag{3.7}$$

The important point to note from the inequalities (3.2) and (3.5) is that the infima over Ω of expressions of the form $P(H_{(k)} \geq K)$ are attained when $\theta_{[1]} = \dots = \theta_{[k]}$.

For the particular case when $a(r) = r$, $nH_i = T_i$, the rank sum statistic associated with π_i. Denoting $R_i(G)$ by R_i in this case, the infimum of $P(CS|R_i)$ can be related to the Mann-Whitney statistic. If U is the Mann-Whitney statistic associated with samples of size n and (k-1)n taken from identically distributed populations, then

$$\inf_{\Omega} P(CS|R_1) = P(U \geq nd). \tag{3.8}$$

A similar expression can be derived for R_2. The Mann-Whitney U-statistic has been tabulated by Milton (1964) among others.

Since $\sum_{j=1}^{k} H_{(j)} = A/n$, we see that

$$\max_{1 \le j \le k} H_j \ge A/nk. \tag{3.9}$$

Hence, a sufficient, but not necessary, condition for the selection rule $R_3(G)$ to select a nonempty subset is that P* be sufficiently large so that

$$D \le A/N. \tag{3.10}$$

For large n, this sufficiency condition for rule $R_3(G)$ is satisfied if $P^* > \frac{1}{2}$. For rule R_3, i.e. when $a(r) = r$, the condition is $D \le (N+1)/2$. As an example, with $k = 3$, $n = 5$ the sufficient condition $D \le 8$ is satisfied for $P^* \ge 0.523$ and for such values a nonempty subset will be selected.

The evaluation of the constants $D = D(k,n,P^*)$ for the rule R_3 can be effected as follows:

$$P^* \le P(T_i \ge Dn) = P(U \le n^2(k-\tfrac{1}{2})-n(D-\tfrac{1}{2})), \tag{3.11}$$

where now we consider all populations identically distributed. Hence, Dn is the largest integer satisfying the inequality (3.10).

4. The Exact and Asymptotic Distribution of $\max_{1 \le j \le k} T_j - T_i$ for Identically Distributed Populations.

In this section the random variables X_{ij}, $j = 1,\ldots,n_i$; $i = 1,\ldots,k$, are assumed independent identically distributed with a continuous distribution $F(x)$. In this case the H_i are exchangeable random variables if $n_i = n$, $i = 1,\ldots,k$. It should be noted that in a slippage-type configuration, the constants required to implement rules $R_i(G)$, $i = 1,2,3$, are determined from the basic probability requirement $P(CS|R_i(G)) \ge P^*$ calculated with identically distributed populations. In the case $a(R_{ij}) = R_{ij}$ the procedures $R_i(G)$ reduce to the rank sum procedures R_i, $i = 1,2,3$. The distribution of the statistic $\max_{1 \le j \le k} T_j - T_1$,

both exact and asymptotic, is somewhat easier to obtain than the corresponding distribution of the statistic $\max_{1\le j\le k} T_j/T_1$. For some results concerning the latter statistic, see McDonald (1969). Our concern here will be the former which is tantamount to considering rule R_1. Corresponding to rule R_3 is the statistic T_1, the distribution of which has been well-treated elsewhere in the Mann-Whitney format.

Gupta and McDonald (1970) have tabulated the quantity $P(T_1 \ge \max_{1\le j\le 3} T_j-m)$ for n = 2(1)5 and $m = 0,1,\ldots,2n^2$ (which covers the entire distribution). Asymptotically (as $n \to \infty$), they show

$$P[T_k \ge \max_{1\le j\le k} T_j-m] \to \int_{-\infty}^{\infty}[\Phi(x+m/z)]^{k-1}\,\varphi(x)dx \quad (m \ge 0), \tag{4.1}$$

where $\Phi(\cdot)$ and $\varphi(\cdot)$ are the cumulative distribution function and density of a standard normal random variable, respectively, and

$$z = z(n,k) = n[k(nk+1)/12]^{1/2}. \tag{4.2}$$

Integrals of the type

$$\int_{-\infty}^{\infty}[\Phi(x+h2^{1/2})]^{k-1}\,\varphi(x)dx = P^* \tag{4.3}$$

have been considered in several publications. The h quantity appearing in this expression has been tabulated (to 3 dp) by Gupta (1963) in Table I for k = 2(1)51 and P* = .75, .90, .95, .975, and .99. Similar values are provided (to 4 dp) in Table 1 of Gupta, Nagel and Panchapakesan (1973) for the same P* and k = 2(1)11(2)51. Additional tabulation of h is provided by Milton (1963). In Table IB of Milton's report, the h quantity is tabulated (to 6 dp) for k = 3(1)10(5)25 and P* = .3(.05).95, .975, .99, .995, .999, .9995, and .9999. In Table II of the same publication P* values are given (to 8 dp) for h = 0(.05)5.15

for all the previously mentioned values of k. Thus, this asymptotic value can be obtained from a variety of sources and can be applied directly to very large data sets - up to 51 populations and any (large) sample size.

Matching the right hand side of (4.1) with (4.3) yields an asymptotic approximation to $m = nd$ given by

$$\tilde{m} = hn[k(nk+1)/6]^{1/2}, \tag{4.4}$$

h being the appropriate solution to (4.3) corresponding to the given P* and k. In the selection rule the smallest integer not less than $\tilde{m}$ should be taken.

Approximations to the Constant m for Use with R_1.

We saw that inf $P(CS|R_1)$ over $\Omega' = \{\underline{\theta}: \theta_{[1]} = \ldots = \theta_{[k-1]} \leq \theta_{[k]}\}$ is attained when $\theta_1 = \ldots = \theta_k$. Suppose we want to evaluate d for which this infimum is at least P*. Using the rank sum statistics, this means that we want the smallest integer $m = nd$ such that

$$P(T_k \geq T_{[k]} - m) \geq P^* \tag{4.5}$$

where the T_i are i.i.d. random variables. McDonald (1971) has discussed two methods of approximating the solution. The first method uses the asymptotic $(n \to \infty)$ expression for the probability given by (4.1).

The second approximation is for large P* (near 1). Suppose $Z_1,\ldots,Z_k$ are N(0,1) random variables with the correlation matrix Σ. Let

$$P\{Z_{[1]} \geq -\delta\} = P^*. \tag{4.6}$$

Dudewicz (1969) has shown that, for large P* (near 1), an approximation to δ is given by

$$\delta^2 \approx -2[\log(1-P^*)] \tag{4.7}$$

in the sense that the ratio tends to 1 as $P^* \to 1$. Using his approximation and the joint asymptotic normal distribution of $[n^2k(nk+1)/6]^{-1/2}(T_k - T_i)$, $i = 1,\ldots,k-1$, we obtain the approximation

$$m^2 \approx -[n^2k \frac{(nk+1)}{3}] \log(1-P^*). \tag{4.8}$$

One can also obtain this approximation from (4.1) by noting that $mz^{-1} \approx \sqrt{2}\,\Phi^{-1}(P^*)$ as $P^* \to 1$, a result of Rizvi and Woodworth (1970), and using the well-known fact that

$$\Phi^{-1}(P^*) \approx [-2 \log(1-P^*)]^{1/2} \quad \text{as} \quad P^* \to 1. \tag{4.9}$$

The two approximations have been compared by McDonald (1971) in the case of $P^* = 0.99$ for $k = 2(1)5$, $n = 5(5)25$.

Let $\hat{m}_1$ and $\hat{m}_2$ denote the approximate values of m given by (4.4) and (4.8), respectively. The numerical evaluations of $\hat{m}_1$ and $\hat{m}_2$ show that (a) $\hat{m}_2 - \hat{m}_1$ increases in n for fixed k, and decreases in k for fixed n, (b) $\hat{m}_1/\hat{m}_2$ increases in k for fixed n, and is constant for fixed k over various values of n, and (c) both approximations are conservative, $\hat{m}_2$ being more so than $\hat{m}_1$. For $k = 2$, McDonald (1971) has analytically shown that $\hat{m}_2 - \hat{m}_1$ is positive and increasing in n, and that $\hat{m}_2/\hat{m}_1$ is independent of n.

5. Comparisons between R_1 and R_3 and with Parametric Procedures.

Recall that for $k = 2$ the rules $R_i(G)$, $i = 1, 2, 3$, are equivalent. For the special case of rank sum statistics based on equal sample sizes, Gupta and McDonald (1970) have studied the asymptotic efficiency of R_1 relative to the means procedure of Gupta (1956) for normal populations and the efficiency of R_2 relative to the procedure of Gupta (1963) for gamma populations both in the case of $k = 2$ populations.

Let π_1 and π_2 be independent normal populations with means θ_0 and $\theta_0+\theta(\theta \geq 0)$ and common unit variance. Let R denote Gupta's means procedure. For both R_1 and R satisfying the P*-condition, the asymptotic efficiency of R_1 relative to R is $ARE(R_1,R;\theta) = \lim_{\epsilon\to 0} n_R(\epsilon)/n_{R_1}(\epsilon)$, where $n_R(\epsilon)$ and $n_{R_1}(\epsilon)$ are the sample sizes for which $E(S)-P(CS)=\epsilon$ for R and R_1, respectively. It is shown by Gupta and McDonald that

$$ARE(R_1,R;\theta) = \{\frac{2\Phi(\theta/\sqrt{2})-1}{2\theta B(\theta)}\}^2, \tag{5.1}$$

where

$$B^2(\theta) = \int_{-\infty}^{\infty} \Phi^2(x+\theta)\varphi(x)dx-\Phi^2(\theta/\sqrt{2}).$$

As θ decreases to zero, we see that $ARE(R_1,R;\theta) \to 3/\pi = 0.9549$.

Some exact calculations for the probabilities of choosing π_1 and π_2 using rank sum procedures can be made using Table C-1 of Milton (1970) for $\theta = .2(.2)1.0$, 1.5, 2.0, and 3.0. This table tabulates the distribution of the Wilcoxon two-sample statistic under the normal shift alternative specified by θ. As an example, for k = 2, n = 6, and P* = .910177, the rank sum selection rules take the form: select π_i iff $T_i \geq 31$, i = 1,2. If the underlying distributions are normal with means 0 and $\theta = .2$ with unit variances, then by summing the appropriate rows in Table C-1 we find $P(T_1 \geq 31)$ = P(Choosing π_1) $\doteq .8465$ and $P(T_2 \geq 31)$ = P(Choosing π_2) $\doteq .9518$.

Let R' denote Gupta's procedure for gamma populations. Let the scale parameters of π_1 and π_2 be θ_0 and $\theta_0\theta$, $\theta > 1$. In this case

$$ARE(R_2,R';\theta) = [\frac{(\theta-1)}{4(\theta+1)B(\theta)\log\theta}]^2, \tag{5.2}$$

where now

$$B^2(\theta) = 1-2(1+\theta)^{-1} + (2\theta+1)^{-1} + \theta(2+\theta)^{-1}-2\theta^2(1+\theta)^{-2}.$$

As θ decreases to 1, we have $ARE(R_2,R';\theta) \to 3/4$.

In another paper Gupta and McDonald (1972), have compared the procedures R_1, R_2, and R_3 based on rank sum statistics with a procedure R_m which they proposed for selection from gamma populations in terms of the guaranteed life. Let π_i *have the associated density function*

$$f(x-\theta_i) = \begin{cases} [\lambda\Gamma(r)]^{-1}[(x-\theta_i)/\lambda]^{r-1}e^{-(x-\theta_i)/\lambda}, & x \geq \theta_i \\ 0 & \text{elsewhere,} \end{cases}$$

where $r(> 0)$ and $\lambda(> 0)$ are known common parameters. In life-testing problems, the parameter θ is called the guaranteed life time. We can assume with no loss of generality that $\lambda = 1$. Let $Y_i = X_{i[1]}$, be the smallest order statistics based on n independent observations from π_i. It is known that Y_i is complete and sufficient statistic for θ_i. The procedure R_m of Gupta and McDonald for selecting a subset containing tne population with the largest guaranteed life is

R_m: Select π_i iff

$$Y_i > Y_{[k]}-b, \tag{5.3}$$

where $b = b(k,n,P^*) > 0$ is chosen to satisfy the P*-requirement. They have shown that

$$\inf_{\Omega} P(CS|R_m) = \int_0^{\infty} H^{k-1}(x+b)dH(x), \tag{5.4}$$

where $H(x)$ is the cdf of Y_i when $\theta_i = 0$.

In the special case of $r = 1$, the exponential case, (5.4) reduces to

$$\inf_{\Omega} P(CS|R_m) = k^{-1}(1-w)^{-1}(1-w^k), \tag{5.5}$$

where $w = 1-e^{-nb}$. For this special case, Gupta and McDonald (1972) have tabulated the values of b for $k = 5, 10$; $n = 2(1)25$; and $P^* = 0.75, 0.90, 0.95, 0.975, 0.99$.

Consider three exponential populations with location parameters $\theta_1 = 0$, $\theta_2 = \theta_3 = \theta \geq 0$. In this case, Gupta and McDonald (1972) have compared the expected subset sizes for the procedures R_1, R_2, R_3, and R_m for $\theta^* = 0(0.1)1.5$ and $P^* = 0.6$, 14/15. The computations indicate that (1) R_1 and R_2 perform equally well for $P^* = 0.6$, (2) R_2 and R_3 perform equally well for $P^* = 14/15$, (3) $E(S|R_2) = E(S|R_3) \leq E(S|R_1)$ for all θ, equality holding for $\theta = 0$, (4) R_m performs better than all the distribution-free procedures for the smaller value of P^*, (5) for the larger P^*, the distribution-free procedures are better than R_m for $\theta \leq 0.5$, and (6) for larger values of $\theta (\theta \geq 1.1) R_m$ is the best among the four rules.

Ofosu (1974) has studied the procedure R_m and compares its performance with a procedure that excludes from the selected subset those populations for which Y_i is sufficiently below $\bar{Y}$, the average of the Y_i. Based on a comparison of the expected subset sizes, Ofosu concludes that R_m is superior to the rules based on averaged Y_i in almost all situations. For those rare situations where R_m is not superior, it is only slightly inferior.

Gupta and McDonald (1970) compare the performance of selection rules R_1 and R_3 in some Monte Carlo studies. Normal and logistic distributions with variance unity were studied for different configurations of their means. For $k = 3$ and $n = 2,3,4$, these configurations were taken to be (0·1,0,0), (0·2,0,0), (0·5,0,0),(1·0,0,0), (2·0,0,0), (0·1,0·1,0), (0·2,0·2,0), (0·5, 0·5,0), (1·0,1·0,0), (2·0,2·0,0). The number of simulations were 500 or 1000. The logistic distribution was chosen because equally spaced scores such as ranks yield locally most powerful tests for the location parameter of this distribution. The constants d and D were chosen to yield approximately the same P^* in the case of identical distributions. Then the ratio of $kP(CS|R)$ and $E(S|R)$ was computed for both rules R_1 and R_3. The bigger ratio for a rule indicates it to be

better than the other. For example, for $k = 3$, $n = 2$, then $D = 2$ and $d = 3$ give the probability 14/15 for the identical case. Using the configuration (0·1,0,0) for the normal means, the two ratios are 1·012 for R_1 and 1·005 for R_3 so that R_1 seems slightly better than R_3. Using the configuration (0·5,0,0), R_3 was slightly better than R_1; the ratios being 1·045 for R_1 and 1·049 for R_3.

These Monte Carlo studies showed no significant uniform superiority of either of these procedures. However, R_3 seemed to perform slightly better than R_1 in the cases where the two highest parameters are equal. No difference in the performance of R_1 and R_3 was noticeable when the distribution changed from logistic to normal. In all cases the frequency of correct selections for R_1 was higher than the theoretical value exactly calculated for the identical distributions. Thus, there was no indication that the infimum of the probability of a correct selection does not take place when all populations are identically distributed as normal or logistic distributions under shift in location.

6. Selection Procedures Based on Pairwise Ranks

As noted earlier the least favorable configuration over Ω for the selection rule $R_1(G)$ is not known and a counterexample exists showing that the infimum of the probability of a correct selection does not occur when all populations are identically distributed (Rules of the form $R_3(G)$ do not share this difficulty). Hsu (1980) overcomes this difficulty by constructing a rule based on pairwise rather than joint ranking of the samples.

Let $R_{j\ell}^{(i)}$ denote the rank of $X_{j\ell}$ within $X_{i1},\ldots,X_{in},X_{j1},\ldots,X_{jn}$, and let $R_j^{(i)} = \sum_{\ell=1}^{n} R_{j\ell}^{(i)}$ be the rank sum statistic for π_j compared to π_i. Let $\{D_\ell^{(ji)}, \ell = 1,\ldots,n^2\}$ denote the collection of n^2 ordered differences $X_{iu}-X_{jv}$, $u, v = 1,\ldots,n$ and set $D_{med}^{(ji)} = \text{median}\{D_\ell^{(ji)}\}$, i.e., $D_{med}^{(ji)}$ is the usual Hodges-Lehmann (H-L) estimator of $\theta_i-\theta_j$. For $i = 1,\ldots,k$, let $M_i = k^{-1}\sum_{j=1}^{k} D_{med}^{(ji)}$, where

$D_{med}^{(ii)} = 0$. The procedure proposed by Hsu for selection of the population with the largest θ_i, denoted by R_R, is as follows:

R_R: Select π_i iff

$$M_i = \max_{1 \le j \le k} M_j \quad \text{or} \quad \max_{j \ne i} R_j^{(i)} < r_n(P^*), \tag{6.1}$$

where $r_n(P^*)$ is the smallest integer such that $P_0[\max_{j \ne i} R_j^{(i)} < r_n(P^*)] \geq P^*$ and $P_0(\cdot)$ indicates the probability is computed assuming all populations are identically distributed.

The procedure R_R does not depend on the pairwise ranks alone. However, the contribution from the "$M_i = \max_{1 \le j \le k} M_j$" portion is small when n is large, and is included to insure that a nonempty subset is selected. The constants $r_n(P^*)$ can be obtained from Steel (1959) for $P^* = .95, .99$, $k = 3(1)10$, $n = 4(1)20$; from Miller (1966, Table VIII) for $P^* = .95, .99$, $k = 3(1)11$, $n = 6(1)20(5)50, 100$.

Hsu also investigates the Pitman efficiency of the R_R procedure compared to a means procedure (with common unknown variance) and shows it to be the same as the Pitman efficiency of the Mann-Whitney test relative to the usual t-test.

Letting $D_{(1)}^{(ji)} \le D_{(2)}^{(ji)} \le \ldots \le D_{(n^2)}^{(ji)}$ denote the n^2 ordered values of $D_\ell^{(ji)}$, $m = r_n(P^*) - n(n+1)/2$, and $D_{(m)}^{(\cdot i)} = k^{-1} \sum_{j \ne i} D_{(m)}^{(ji)}$, an alternative procedure was also suggested by Hsu and is given by:

R_R': Select π_i iff

$$\min_{j \ne i}(D_{(m)}^{(\cdot i)} - M_j) > 0 \quad \text{or} \quad \max_{j \ne i} R_j^{(i)} < r_n(P^*). \tag{6.2}$$

The subset selected by R_R' always contains the subset selected by R_R; however, the two rules are equi-efficient in terms of their Pitman efficiencies under the location model.

7. Selection Procedures Based on Vector Ranks.

In the preceding procedures R_1, R_2, and R_3, the statistics H_i are defined using the ranks of the observations in the pooled sample. In cases with equal sample sizes, vector-at-a-time sampling can be used effectively to remove block effects, such as in a two-way layout, and to reduce data storage requirements. These procedures cover, for example, models of the form

$$X_{ij} = \mu + \theta_i + \beta_j + \epsilon_{ij}, \quad i = 1,\dots,k, \quad j = 1,\dots,n, \tag{7.1}$$

where θ refers to a population effect, β to a block effect, and ϵ to an error term with any (not necessarily normal) continuous distribution.

Let $(X_{1j}, X_{2j}, \dots, X_{kj})$ be the jth vector and R_{ij} be the rank of X_{ij} among the k observations of the vector. Let $Z(1) \le Z(2) \le \dots \le Z(k)$ denote an ordered sample of size k from a continuous distribution G. Define $a(r)$ as in Section 2, i.e.

$$a(r) = E[Z(r)|G], \qquad r = 1,\dots,k,$$

and set

$$J_i = n^{-1} \sum_{j=1}^{n} a(R_{ij}), \quad i = 1,\dots,k. \tag{7.2}$$

McDonald (1972) investigated the classes of procedures $R_1'(h;G)$ and $R_2'(g;G)$ which are defined using the two classes of functions $\{h(x)\}$ and $\{g(x)\}$, where h and g are nondecreasing real-valued functions defined on the interval $I = [b(1), b(k)]$ and h satisfies the additional property that $h(x) \ge x$ for all $x \in I$. The two classes of procedures are

$R_1'(h;G)$: Select π_i iff

$$h(J_i) \ge J_{[k]}, \tag{7.3}$$

and

$R_2'(g;G)$: Select π_i iff

$$g(J_i) \ge 0. \tag{7.4}$$

Particular members of these classes that are of special interest are $R_1'(G)$ with $h(x) = x+b$, $b > 0$, and $R_2'(G)$ with $g(x) = x-d$, d real. Of course, $R_2'(g;G)$ can select an empty set; however, the rule $R_2(G)$ will necessarily choose a nonvoid subset if $P^* > .5$ and n is large. The treatment of $R_1'(G)$ and $R_2'(G)$ parallels that of $R_1(G)$ and $R_2(G)$ described earlier. The infimum of $P(CS)$ over Ω is attained at a point in Ω_k in the case of $R_1'(h;G)$. However as in the case of $R_1(G)$, it is not generally true that the infimum is attained at an equi-parameter configuration. But the statement is true in the case of $R_2'(g;G)$.

When $b(r) = r$, $nH_i = T_i$, the rank sum statistic associated with π_i. McDonald (1973) has discussed the related distribution of $U = \max_{1 \leq j \leq k} T_j - T_1$, where the distributions F_i are identical. He has tabulated $P(U \leq b)$ for $k = 2$, $n = 2(1)20$; $k = 3$, $n = 2(1)8$; $k = 4$, $n = 2(1)5$; and $k = 5$, $n = 2,3$. For $P(U \leq b) = P^* = 0.75, 0.90, 0.95, 0.975$, and 0.99, he has tabulated the asymptotic value of b for $k = 2$, $n = 10(5)20$; $k = 3$, $n = 6(1)8$; $k = 4$, $n = 3(1)5$; and $k = 5$, $n = 3$.

The investigations of McDonald (1973) with respect to slippage configuration based on simulations show that R_1' and R_2' (which are $R_1'(G)$ and $R_2'(G)$, respectively, in the special case with $b(r) = r$) are roughly equivalent when the underlying distribution has a long tail and the slippage is small, and that R_1 is better otherwise. These rules have been used by McDonald (1979) in an analysis of state traffic fatality rates recorded by year.

Lorenzen and McDonald (1980) further investigate the probability of a correct selection using rule R_1' by Monte Carlo simulations covering a wide range of distributions and parameter configurations (both location and scale). In all cases investigated the LFC, i.e., the configuration minimizing $P(CS)$, appeared to be the equi-parameter configuration. This suggests that the practical inference corresponding to the selection procedure need not be restricted to the slippage configurations.

In another paper, McDonald (1975) considered the case of three exponential distributions with parameters (guaranteed lives) $0 = \theta_1 \leq \theta_2 = \theta_3 = \theta$ with samples of size two. For the rules R_1, R_2, and R_3 (using rank sum statistics T_i) the infimum of P(CS) takes place when $\theta_1 = \theta_2 = \theta_3$. However, it is shown that the expected subset size is not bounded above by kP*, a property enjoyed by many parametric procedures [see Gupta (1965)] for the location parameter case under monotone likelihood ratio conditions.

Within the context of a block design (2-way classification) Lee (1980) considers another type of selection rule based on the statistics $Y_i = \sum_{j=1}^{n} Y_{ij}$, $i = 1,\ldots,k$, where

$$Y_{ij} = \begin{cases} 1, & \text{if } X_{ij} = \max\limits_{1 \leq \ell \leq k} X_{\ell j} \\ 0, & \text{otherwise.} \end{cases} \tag{7.5}$$

The selection rule is stated as

R_{MS}: Select π_i iff

$$Y_i \geq \max_{1 \leq j \leq k} Y_j - d_{MS}, \tag{7.6}$$

where d_{MS} is the smallest nonnegative integer required to insure the probability of a correct selection is no less than a prescribed P*. The procedure is a multinomial selection rule (hence the subscript MS) designed to choose a subset to contain the population having the highest probability of yielding the largest observation. An analogous rule for choosing a subset to contain the population having the highest probability of yielding the smallest observation is also defined.

The constants d_{MS} required to implement the procedure R_{MS} have been determined by Lee (1980) using Monte Carlo simulation, assuming the underlying distributions are identical, for $k = 49$ and $n = 17$. These values were then used to select subsets of states on the basis of traffic fatality rates recorded over a period of 17 years. Gupta and Nagel (1967) investigated the least favorable configuration in a corresponding multinomial formulation and concluded, based on some numerical case studies, that the identically distributed case appears least favorable. Panchapakesan (1971) proved that the identically distributed configuration is asymptotically least favorable.

8. Selection Procedures Based on Hodges-Lehmann Estimators

Let X_{ij} ($j = 1,\ldots,n$; $i = 1,2,\ldots,k$), $k \geq 2$, be independent random observations from k populations with continuous cdf's $F(x-\theta_i)$, $i = 1,2,\ldots,k$, with common variance $\sigma^2 = 1$. The following problems have been considered by Bechhofer (1954) under the normality assumption:

(i) Select a "good" population, the ith population being regarded as good if $\theta_i \geq \theta_{[k]}-\Delta$, for some preassigned $\Delta > 0$ ($i = 1,2,\ldots,k-1$);

(ii) select the best t populations, i.e., the populations with location parameters $\theta_{[k-t+1]},\ldots,\theta_{[k]}$ without regard to order;

(iii) select the best t populations with regard to order.

His approach, now known as the "indifference zone" approach selects the "best" populations with a guaranteed minimum probability P^* (preassigned) of correct selection when $(\theta_1,\ldots,\theta_k)$ lies in a subset, say Ω' of the parameter space. The region Ω' is called the preference zone and $R^k-\Omega'$ is the indifference zone. Some of the procedures discussed earlier use rank statistics for selection purposes.

However, when formulated for the problems discussed in this section, the slippage configuration of parameters defined by the indifference zone is not necessarily the LFC. The slippage configuration as pointed out by Puri and Puri (1969) is least favorable when the parameters satisfy the relation $\theta_{[i]} - \theta_{[j]} = o(n^{-\frac{1}{2}})$ for all $1 \leq i, j \leq k$, $i \neq j$.

Ghosh (1973) has proposed alternate procedures, based on one-sample Hodges-Lehmann estimators of θ_i's under the additional assumption that F is symmetric about the origin. Ghosh's procedures give in all these cases least favorable configurations for finite n without needing any restriction on the parameters.

Gupta and Huang (1974) have proposed some procedures to select a subset of the given k populations which is guaranteed to exclude all bad populations with probability not less than some preassigned P*.

Let $R_{ij} = \frac{1}{2} + \sum_{\ell=1}^{n} u(|X_{ij}|-|X_{i\ell}|)$, $j = 1,2,\ldots,n$, $i = 1,2,\ldots,k$, where $u(t) = 1, \frac{1}{2}$, or 0 at $t >, =$, or < 0. Thus R_{ij} is the rank of $|X_{ij}|$ among $|X_{i1}|,\ldots,|X_{in}|$ $(1 \leq i \leq k;\ 1 \leq j \leq n)$. Let $\underline{X}_i' = (X_{i1},\ldots,X_{in})$. Consider the one-sample signed rank statistics

$$h(\underline{X}_i) = \sum_{j=1}^{n} \operatorname{sgn}(X_{ij}) EJ(U_{nR_{ij}}), \tag{8.1}$$

$i = 1,2,\ldots,k$, where $\operatorname{sgn}(t) = 1, 0$ or -1 according as $t >, =$, or < 0; $U_{n1} \leq U_{n2} \leq \cdots \leq U_{nn}$ are the n ordered random variables from a rectangular (0,1) distribution, and $J(u) = \psi^{-1}(\frac{1+u}{2})$, where $\psi(x)$ is the df of a random variable satisfying $\psi(x) + \psi(-x) = 1$ for all real x.

The one-sample H-L estimators are given by

$$\hat{\theta}_i(\underline{X}_i) = \frac{1}{2} \{\hat{\theta}_{i1}(\underline{X}_i) + \hat{\theta}_{i2}(\underline{X}_i)\}, \tag{8.2}$$

$i = 1,2,\ldots,k$, where $\hat{\theta}_{i1} = \sup\{a: \ h(\underline{X}_i - a\underline{1}_n) > 0\}$, $\hat{\theta}_{i2}(\underline{X}_i) = \inf\{a: h(\underline{X}_i - a\underline{1}_n) > 0\}$, $\underline{1}_n' = (1,\ldots,1)$ is an n-tuple with all elements 1. All these statistics and estimators depend on n. The following property of location invariance (see Hodges and Lehmann (1963)) is satisfied by these estimators:

$$\hat{\theta}_i(\underline{X}_i + c\underline{1}_n) = \hat{\theta}_i(\underline{X}_i) + c, \tag{8.3}$$

$i = 1,2,\ldots,k$, c being any constant. In the particular case when $J(u) = u$ or $\chi_1^{-1}(u)$ (the inverse of a chi-distribution with one degree of freedom) the statistics become the Wilcoxon signed-rank or normal-score statistics. In the former case

$$\hat{\theta}_i(\underline{X}_i) = \underset{1 \le j \le j' \le n}{\text{med}} \frac{X_{ij} + X_{ij'}}{2}, \quad i = 1,2,\ldots,k.$$

Let $\hat{\theta}_{[1]} \le \hat{\theta}_{[2]} \le \cdots \le \hat{\theta}_{[k]}$ denote the ordered estimators and let $\hat{\theta}_{(i)}$ be the unknown estimator associated with $\theta_{[i]}$ $(1 \le i \le k)$.

An Elimination Type Procedure to Select a Subset Excluding All "Strictly Non t Best" Populations

Let $d(\theta_i, \theta_j)$ be a suitable distance measure between θ_i and θ_j; the population π_i is "strictly non t best" if $d(\theta_{[k-t+1]}, \theta_i) = \theta_{[k-t+1]} - \theta_i > \Delta$, where Δ is a given positive constant. Let m denote the unknown number of "strictly non t best" populations in the given collection of k populations. Clearly, we have $0 \le m \le k-t$. Let $\Omega_m = \{\underline{\theta}: \theta_{[1]} \le \cdots \le \theta_{[m]} < \theta_{[k-t+1]} - \Delta \le \theta_{[m+1]} \le \cdots$
$\le \theta_{[k-t+1]} \le \cdots \le \theta_{[k]}\}$.

Then

$$\Omega = \bigcup_{m=0}^{k-t} \Omega_m.$$

Let CD stand for a correct decision, which is defined to be the selection of a subset which excludes all the "strictly non t best" populations. Gupta and Huang (1974) define the rule R as follows:

R: Reject π_i iff

$$\hat{\theta}_i < \hat{\theta}_{[k-t+1]} - \Delta + d_1 \quad (0 < d_1 < \Delta). \tag{8.4}$$

The constant d_1 is chosen to be the smallest number such that

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CD|R) \geq P^*.$$

Gupta and Huang (1974) have shown that $P_{\underline{\theta}}(CD|R)$ is a nonincreasing function of $\theta_{[i]}$ $(i = 1,\ldots,m)$ and a nondecreasing function of $\theta_{[i]}$ $(i = m+1,\ldots,k)$. Hence

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CD|R) = \inf_{0\leq m\leq k-t} \inf_{\underline{\theta}\in\Omega_m} P_{\underline{\theta}}(CD|R).$$

It is known that if θ_i's are true values of the parameters, then under some regularity assumptions $\sqrt{n}\,(\hat{\theta}_i(\underline{X}_i)-\theta_i)\,B(F)/A$ tends asymptotically (as $n \to \infty$) to Y_i with $N(0,1)$ where $A^2 = \frac{1}{4}\int_0^1 J^2(u)du$, $B(F) = \int_0^\infty \frac{d}{dx} J(2F(x)-1)dF(x)$. These statistics Y_i's are mutually independent. This leads to a lower bound on the infimum of the probability of a correct decision for large n as follows:

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CD|R) \geq \frac{t!}{r!(t-r-1)!}\int_{-\infty}^{\infty} \Phi^{k-t}(x+d\sqrt{n})\Phi^{r}(x)[1-\Phi(x)]^{t-r-1}\varphi(x)dx, \tag{8.5}$$

where $r=\min(t, k-t-1)$, $d = \sqrt{0.864}\, d_1$ (or d_1) for the Wilcoxon (or normal score) case. For the case $F(x) = \Phi(x)$, then using normal scores the inequality (8.5) is an equality and the result agrees with that obtained by Carroll, Gupta and Huang (1976).

It has also been shown that

$$\lim_{n\to\infty} \inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CD|R(n))$$

$$= \lim_{n\to\infty} \frac{t}{r'(t-r-1)!} \cdot \int_{-\infty}^{\infty} \Phi^{k-t}\left(x + \frac{B(F)}{A} d_1\sqrt{n}\right)\Phi^{r}(x)[1-\Phi(x)]^{t-r-1}\varphi(x)dx$$

$$= 1, \text{ since } \frac{B(F)}{A} d_1 > 0 ,$$

so that the sequence of rules $\{R(n)\}$ is consistent wrt Ω.

Since the cdf of each $\hat{\theta}_i(\underline{X}_i)$ is stochastically nondecreasing in θ_i, it follows that for every $\underline{\theta} \in \Omega$ and $1 \leq i < j \leq k$.

$$P_{\underline{\theta}}\{R(n) \text{ rejects } \pi_{(i)}\} \geq P_{\underline{\theta}}\{R(n) \text{ rejects } \pi_{(j)}\},$$

and thus $R(n)$ is a so-called monotone procedure.

References

Bartlett, N. S. and Govindarajulu, Z. (1968). Some distribution-free statistics and their application to the selection problem. Ann. Inst. Statist. Math., 20, 79-97.

Bechhofer, R. E. (1954). A single sample decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist., 25, 16-39.

Carroll, R. J., Gupta, S. S. and Huang, D. Y. (1976). On selection procedures for the t best populations and some related problems. Comm. Statist., 4, 987-1008.

Dudewicz, E. J. (1969). An approximation to the sample size in selection problems. Ann. Math. Statist., 40, 492-497.

Ghosh, M. (1973). Nonparametric selection procedures for symmetric location parameter populations. Ann. of Statist. 1, 773-779.

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Tech. Report No. 150, Institute of Statistics, University of North Carolina.

Gupta, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math., 14, 199-216.

Gupta, S. S. (1963). Probability integrals of multivariate normal and multivariate t. Ann. Math. Statist., 34, 792-828.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics, 7, 225-245.

Gupta, S. S. and Huang, D. Y. (1974). Nonparametric subset selection procedures for the t best populations. Bull. Inst. Math., Academia Sinica, 2, 377-386.

Gupta, S. S. and McDonald, G. C. (1970). On some classes of selection procedures based on ranks. Nonparametric Techniques in Statistical Inference. (Editor M. L. Puri), Cambridge University Press, Cambridge, England, 491-514.

Gupta, S. S. and McDonald, G. C. (1972). Some selection procedures with applications to reliability problems. Operations Research and Reliability (Proceedings of the NATO Conference, Italy, 1969). Gordon and Breach, New York, 421-439.

Gupta, S. S. and Nagel, K. (1967). On selecting and ranking procedures and order statistics from the multinomial distribution. Sankhyā, B, 29, 1-34.

Gupta, S. S., Nagel, K., and Panchapakesan, S. (1973). On the order statistics from equally correlated normal random variables. Biometrika, 60, 403-413.

Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. John Wiley, New York.

Hodges, J. L. Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. Ann. Math. Statist., 34, 598-611.

Hollander, M. and Wolfe, D. A. (1973). Nonparametric Statistical Methods. John Wiley, New York.

Hsu, J. C. (1980). Robust and nonparametric subset selection procedures. To appear: Commun. Statist.-Theor. Meth.

Lee, Y. J. (1980). Nonparametric selections in blocked data: application to motor vehicle fatality rate data. Submitted for publication.

Lehmann, E. L. (1975). Nonparametrics: Statistical Methods Based on Ranks. McGraw-Hill, New York.

Lorenzen, T. J. and McDonald, G. C. (1980). A nonparametric analysis of urban and rural traffic fatality rates. General Motors Research Publication, General Motors Research Laboratories, Warren, Michigan.

McDonald, G. C. (1969). On some distribution-free ranking and selection procedures. Department of Statistics, Purdue University. Mimeograph Series No. 174.

McDonald, G. C. (1971). On approximating constants required to implement a selection procedure based on ranks. Statistical Decision Theory and Related Topics.(Eds. S. S. Gupta and J. Yackel), Academic Press, New York, pp. 299-312.

McDonald, G. C. (1972). Some multiple comparison selection procedures based on ranks. Sankhyā, A, 34, 53-64.

McDonald, G. C. (1973). The distribution of some rank statistics with applications in block design selection problems. Sankhyā, A, 35, 187-204.

McDonald, G. C. (1975). Characteristics of three selection rules based on ranks in the 3x2 exponential case. Sankhyā, B, 36, 261-266.

McDonald, G. C. (1979). Nonparametric selection procedures applied to state traffic fatality rates. Technometrics, 21, 515-523.

Miller, R. G. (1966). Simultaneous Statistical Inference. McGraw-Hill, New York.

Milton, R. C. (1963). Tables of the equally correlated multivariate normal probability integral. Tech. Report No. 27, Department of Statistics, University of Minnesota.

Milton, R. C. (1964). An extended table of critical values for the Mann-Whitney (Wilcoxon) two-sample statistic. J. Am. Statist. Assoc., 59, 925-34.

Milton, R. C. (1970). Rank Order Probabilities: Two-Sample Normal Shift Alternatives. John Wiley, New York.

Ofosu, J. B. (1974). On selection procedures for exponential distributions. Bull. Math. Statist., 16, 1-9.

Panchapakesan, S. (1971). On a subset selection procedure for the most probable event in a multinomial distribution. Statistical Decision Theory and Related Topics. (Eds. S. S. Gupta and J. Yackel), Academic Press, New York, 275-298.

Puri, M. L. and Puri, P. S. (1969). Multiple decision procedures based on ranks for certain problems in analysis of variance. Ann. Math. Statist. 40, 619-632.

Rizvi, M. H. and Woodworth, G. G. (1970). On selection procedures based on ranks: Counterexamples concerning least favorable configurations. Ann. Math. Statist., 41, 1942-1951.

Steel, R. G. D. (1959). A multiple comparison rank sum test: Treatment versus control. J. Am. Statist. Assoc., 54, 767-775.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Mimeograph Series #80-13	2. GOVT ACCESSION NO. AD-A096 093	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Nonparametric Procedures in Multiple Decisions (Ranking and Selection Procedures)		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Mimeo. Series #80-13
7. AUTHOR(s) S. S. Gupta and G. C. McDonald		8. CONTRACT OR GRANT NUMBER(s) ONR N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE May 1980
		13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonparametric procedures, ranking and selection, least favorable configuration, subset selection, rank sum procedures, Mann-Whitney (Wilcoxon) statistic, Hodges-Lehmann estimators elimination-type procedures.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This article surveys statistical techniques which are nonparametric in nature and used in formal ranking and selection of populations. Such methods have been developed only within the last fifteen years and are usually based on rank scores and/or robust estimators (such as the Hodges-Lehmann estimator). The procedures surveyed are applicable to one-way classifications, two-way classifications, and paired-comparison models. Computational methods, useful inequalities, and appropriate numerical tables required to implement these techniques are identified and discussed. Asymptotic relative efficiencies of

the nonparametric methods, compared to their parametric counterparts, are presented. Specific applications of these methods (such as traffic fatality rates) are mentioned and areas for further theoretical and computational research are identified.